Conference Call 5/10/01

The agenda for this call followed an order of priority of items requiring decision.

User Pops

Unduplication Process: Karen's proposal for conducting an inter-rater reliability analysis presents a way to generate at least a degree of pre-process agreement on unduplication parameters among stat officers using sample data. She provided a discussion of the key concepts for understanding probabilistic unduplication as implemented in Automatch just for the DQA team initially, but the discussion paper will be offered to the stat officers after their decision process with the sample data. Team discussion focused on which way to obtain stat officer input: 1)have them identify duplicates among samples of patient data then determine an algorithm that most matches the method(s) they used, or 2) send them the rules we want to use and ask for their input.

Two additional points were made: First, for now, we want to keep unduplication simple, not necessarily in terms of the number of fields used but in the statistical manner in which we process field values. Second, using the first method to obtain stat officer input requires that we be careful to not attach too much meaning to those who are non-responsive.

Team decision: We will send the patient lists to the statistical officers asking them to determine duplicates using their own methods and to include in their responses to us what additional field information would have been helpful during their process. Patricia, Karen, and Stephanie will meet in the meantime to resolve what method is to be used in performing the unduplication process. Target date is end of May for receiving responses from the stat officers and for reaching a decision on the method to use, to allow for producing reports using the agreed upon method by the IHLC meeting at the end of June.

Data Warehouse

Steve Cary (IBM) discussed his initial impressions concerning the need for the establishment of a data warehouse, its features, implementation plans, etc.

Documentation

User Pop Process Logic: A draft will be distributed to the DQA Team today for review.

Workload Reporting Process Logic: A draft of processing logic with reportable criteria will be prepared on major update programs by the end of next week for DQA Team review. When we run into an issue that drills down deeply, we need to draw a line, include the issue on the data issues log, and move on to get the draft out – then return to those issues later.

Access to NPIRS DB

Database Access and Use Policy: The Database Access and Use Policy document is still internal, with initial release estimated on 5/21. Debra describes it as simple but including

guidelines from other agencies. As access expands to other users beyond the HQ Epi program and Area Stat Officers, this policy will also have to evolve.

Connectivity Accomplished! Debra is now connected with DB2 Tools to *testdb*, can see the tables, and has done a simple query. Mike estimates he can get SAS connectivity and tools to Debra tomorrow and then will work on Patricia's access. Debra will help get Linda and Edna connected.

Inventory

Gaps in 97/98 Data: Decision was to try to identify the sites most affected and ask Edna, Linda, Debra, and Patricia to look at the data and recommend either a targeted re-export or extrapolate data to insert from other years (linear correction). NPIRS (Yolinda) has produced a report showing data counts from 1995-2000. The group above will look at this new report, talk with Stephanie, and make a recommendation next week for addressing these gaps.

Workload FY98 and FY99: Review of this data was placed hold. Stephanie hopes to have query results out by end of today or tomorrow. Stan will send her analysis out to all stat officers and we'll place it on the web.

This problem also corrupted values in other fields which may increase duplication in workload data and which requires removal and reloading of the affected visit records. For visit data already approved, we could add a field in the database to store the previous workload reportable flag and within a week have the field added and the reload completed. Then any reports drawn from 1994-97 will use the value in that new field and counts should reflect what was reported previously for those years. Any records for those years that we receive after the field is added will not be counted, but they were not counted previously either. Based on the NPIRS staff recommendation to do this, the decision is to proceed. Open issue: How do we deal with a new export of a record already flagged in this field as reportable – do we overwrite the record all except for the flag field – especially if it's already approved workload.

Other Notes

Joan will respond to webboard messages, send them out to team members, and make sure we respond so that the situation with the message posted by Ross Leatham is not repeated.

Patricia will send Linda's minutes from the recent Stat Officer Workgroup meeting to the team.

Stephanie will work with Karen to develop a spreadsheet for the weekly NPIRS status report so that it includes what has not been done (e.g. export files not yet processed, etc.).

Stephanie will assign tasks to appropriate NPIRS people to review issues and recommend fixes that include an estimated level of effort. NPIRS is also close to getting baseline workload and user pop reports out (data has been processed).

One additional note: We need a task order for Lori Butcher who has been asked to document the PCC export logic.